Thesis for the degree of Doctor of Philosophy

New methods and applications for interferometric GNSS reflectometry

Joakim Strandberg

Department of Space, Earth and Environment
Onsala Space Observatory
Chalmers University of Technology
Gothenburg, Sweden 2020
New methods and applications for interferometric GNSS reflectometry
JOAKIM STRANDBERG
ISBN 978-91-7905-298-0

© JOAKIM STRANDBERG, 2020

Doktorsavhandlingar vid Chalmers tekniska högskola
Ny serie nr. 4765
ISSN 0346-718X

Department of Space, Earth and Environment
Onsala Space Observatory
Chalmers University of Technology
SE-412 96 Gothenburg
Sweden
Telephone: +46 (0)31-772 1000

Cover:

Chalmers Reproservice
Gothenburg, Sweden 2020
New methods and applications for interferometric GNSS reflectometry
JOAKIM STRANDBERG
Department of Space, Earth and Environment
Chalmers University of Technology

Abstract

The GNSS reflectometry technique has been proven to be usable for measuring several environmental properties, such as soil moisture, snow depth, vegetation, and sea level. As numerous GNSS installations are already installed around the world for geodetic purposes, the technique opens up a large data set for new analyses, complementing other environmental measurement campaigns. However, a main drawback of the technique is that its precision generally is worse than more specialised equipment, and while this is in part compensated for its low cost and maintenance requirements, improved precision is still a main goal of research in the field of GNSS reflectometry.

The first topic of this thesis concerns the development of new methods for analysing GNSS-R data to retrieve precise measurements, especially in the case of sea level. As GNSS-R measurements are usually done over time spans of around half an hour, the dynamic sea surface has proven to be a challenge to measure. However, using inverse modelling with least squares adjustment, we prove that we can significantly improve the retrieval precision. Developing on the inverse modelling approach, we also prove that high-precision real-time GNSS reflectometry is also feasible using Kalman filtering.

The other main topic of this thesis is finding new applications for the GNSS-R technique. Firstly, we show that when a GNSS-R installation is mounted close to a body of water, it is possible to determine whether the surface is frozen or not. Secondly, while GNSS reflectometry is traditionally performed with high-precision geodetic instruments, we show that everyday devices, such as a mobile phone, can be used instead. We find that the precision of the mobile devices is on a similar level as for geodetic equipment.

Finally, this thesis explores and highlights one of the challenges that are still left in GNSS-R research: absolute referencing of sea level measurements. Past research has mostly focused on precision, leaving out accuracy, and we show that there are unknown effects that cause an offset between GNSS-R measurements and co-located tide gauges.

Keywords: GNSS, reflectometry, sea level, sea ice
Publications

This thesis is based on the work contained in the following papers:


Other research contributions not included in the thesis:


- D. J. Purnell et al. (2020). Quantifying the Uncertainty in Ground-Based GNSS-Reflectometry Sea Level Measurements. *submitted to IEEE JSTARS*
Acknowledgements

As I write this, the world is facing a pandemic with widespread consequences. And while it has at times brought out much solidarity, with many people doing their part to stop the spread of the virus, it has also brought along a lot of solitariness. A common saying is that "you don’t know what you have until it’s gone", and working from home these past few months has really made me appreciate all the social interactions of a normal working life. Therefore, I can sincerely express a deep thankfulness to all my colleagues at Chalmers who have brightened these five years with interesting discussions over both coffee and lunch. Whether you work in my group, in the same department, or I’ve met you through the Doctoral Student Guild or SACO, I am most grateful of all the times we’ve met.

In times as these, it becomes evident how much we depend on social contacts, and therefore I would also like to offer my warmest gratitude to my family and all my friends, hoping that we all soon can meet together again. I am also especially thankful for my fiancée, Ida, for filling my life with joy. Together, the world is a little less lonely.

Finally, there are two persons without whom this thesis would never have been written: my supervisors, Thomas and Rüdiger. Thank you for all your support, ideas, input, and all your time. And most of all, thank you for believing in me and giving me this chance.

Joakim
# Contents

Abstract .................................................... i  
Publications .............................................. iii  
Acknowledgements ....................................... v  
Contents .................................................. vii  

<table>
<thead>
<tr>
<th>1  Introduction</th>
<th>1</th>
<th>1.1 Interferometric GNSS reflectometry</th>
<th>2</th>
<th>1.2 The Onsala test installation</th>
<th>3</th>
<th>1.3 Thesis overview</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>2  A guide to interferometric GNSS reflectometry</td>
<td>5</td>
<td>2.1 Multipath signals in GNSS</td>
<td>5</td>
<td>2.2 Lomb-Scargle retrievals</td>
<td>8</td>
<td>2.3 Lomb-Scargle in practice</td>
<td>10</td>
</tr>
<tr>
<td>3  New techniques</td>
<td>11</td>
<td>3.1 Inverse modelling of GNSS reflectometry</td>
<td>11</td>
<td>3.2 Kalman filtering for GNSS reflectometry</td>
<td>16</td>
<td>3.3 Particle filter with Lomb-Scargle periodograms</td>
<td>20</td>
</tr>
<tr>
<td>4  New applications</td>
<td>25</td>
<td>4.1 Detecting sea ice</td>
<td>25</td>
<td>4.2 Mobile GNSS reflectometry</td>
<td>26</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5  Atmospheric delay and absolute leveling</td>
<td>29</td>
<td>5.1 Effect of the atmosphere</td>
<td>29</td>
<td>5.2 Applying the corrections</td>
<td>32</td>
<td>5.3 Absolute leveling of GNSS reflectometry</td>
<td>34</td>
</tr>
<tr>
<td>6  Summary of appended papers</td>
<td>37</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Appendix A Derivations 39  
Paper I 49  
Paper II 63  
Paper III 71  
Paper IV 85
Chapter 1
Introduction

Our world is full of noise. From the constant sound of a city to interference on a radio channel, different types of noise are always around us. Mostly, noise is considered a nuisance and something to be avoided, as it disturbs our ability to discern other, more interesting, signals. The sound of the city may drown out the sound of our nature, or the interference on a radio channel may make it impossible to hear the news. However, if the noise itself is analysed it may very well carry some information which can be interesting in its own right, such as the intensity of the traffic. The technique discussed in this thesis is another example of such a case, where something that is usually considered to be noise becomes the signal.

Global Navigation Satellite Systems (GNSS) is a collective term for all satellite systems that are used for positioning and navigation. The most famous is the American GPS constellation, but more systems exist: the Russian GLONASS, the Chinese BeiDou, and the European Galileo constellation are all globally available satellite systems that send out signals which can be used to determine a position anywhere on earth. In ordinary usage of GNSS, it is desirable that the signal travels directly from the satellite to the GNSS device on earth, but sometimes the signal is reflected on one or more surfaces before it reaches the antenna. These reflections are called multipath signals — as they can come from many different directions at once — and are a major source of error in positioning applications (Georgiadou and Kleusberg, 1988). However, in GNSS reflectometry (GNSS-R) the multipath effect is a valuable signal rather than noise.

Multipath signals originate from reflections in the environment around a GNSS antenna, therefore, by analysing the exact impact of the multipath it is possible to retrieve information about the reflecting surfaces (Nievinski and Larson, 2014a). The reflection is affected by various surface properties, and GNSS reflectometry can therefore be used for measuring for example sea level, snow height, soil moisture, and vegetation (Larson et al., 2009, 2008b; Martin-Neira et al., 2001; Rodriguez-Alvarez et al., 2011; Small et al., 2010).
2 Introduction

Figure 1.1: GNSS reflectometry is a close-range remote sensing technique; while the technique measures the immediate surroundings of an antenna, no direct contact with the surface is required. This allows measurements in areas which are not easily directly accessible.

While measurements with GNSS reflectometry are usually less precise than dedicated equipment, they have other benefits. In contrast to many other in situ measurements, GNSS-R measurements typically cover a large area around the antenna (Nievin, 2013), essentially a midway between point measurements and regional measurements from e.g. satellites or airplanes. This also allows the equipment to be mounted at some distance away from the area of interest, in case direct access is unfeasible. Furthermore, having no direct contact can also decrease wear, and thus make the equipment require less maintenance, decreasing the operating costs.

Still, higher precision is always sought after. Therefore, this thesis is dedicated to new and improved methods for GNSS reflectometry, and in the coming chapters I will introduce both new GNSS-R techniques as well as new usage areas.

1.1 Interferometric GNSS reflectometry

The field of GNSS reflectometry has since its inception by Martin-Neira (1993) branched out to several sub fields. The initial concept was developed with spaceborne applications in mind, with ground based installations used only for test purposes (Martin-Neira et al., 2001). However, since then, ground based GNSS reflectometry has become an interesting technique in
The main focus of this thesis is the technique which has come to be called interferometric GNSS reflectometry. The main benefit of this technique is that it can use off-the-shelf equipment to perform environmental measurements, requiring no purposefully built antenna or a special receiver. This means that already installed GNSS antennas can be used, as no specialised hardware is needed; any GNSS installation that happens to be in a location with a view of an interesting area can be used. Standardised hardware also means that the costs for dedicated GNSS-R installations can be kept down.

1.2 The Onsala test installation

In developing the new techniques, much of the testing has been performed on the dedicated GNSS-R installation GTGU at the Onsala Space Observatory, which is in many senses an ideal GNSS-R installation. The antenna is mounted on a beam over the sea surface, which gives a clear view of the sea surface as seen in Figure 1.2. Also, the islets in the inlet of the bay ensure that the waves in the bay are mostly quite small, resulting in a flat and highly reflective surface. Close to the antenna there is also a high precision tide gauge available (Elgered et al., 2019), allowing for accurate comparisons.

As displayed in Figure 1.2, the installation consists of two antennas
mounted in opposite directions. The two antennas can be used together for measuring sea level (Löfgren et al., 2011), but in this thesis only the upward antenna is used. This ensures that the technique is widely implementable without any special installation requirements.

1.3 Thesis overview

The rest of this thesis will go deeper into the field of interferometric GNSS reflectometry, and for most of it I will leave out ‘interferometric’ for brevity. The next chapter, Chapter 2, will focus on previous development of GNSS reflectometry and set the background for my own work. It is meant as a guide for new prospective users who want to learn more about the technique. Then Chapter 3 will continue with introducing a few new methods for GNSS-R retrievals that we have developed, mostly based on Papers I and III on inverse modelling and Kalman filtering, respectively, but also including some previously unpublished experiments with particle filtering. The 4th chapter is based on Papers II and IV, in which we introduce two new ways of using GNSS reflectometry, first for detecting sea ice, and then for using mobile phones and tablets for GNSS-R measurements. In Chapter 5 I discuss some ongoing developments on referencing GNSS-R measurements to an absolute reference point. Finally, Chapter 6 summarises the appended papers and my contributions to them.
Chapter 2
A guide to interferometric GNSS reflectometry

In interferometric GNSS reflectometry, the basic observation is the signal-to-noise ratio (SNR) recorded by GNSS receivers as they track the currently visible satellites. The SNR is affected by several factors, such as antenna design, signal strength, and receiver quality, but most importantly for GNSS reflectometry, it is also affected by multipath signals (Georgiadou and Kleusberg, 1988; Ray and Cannon, 2001). Multipath is usually considered noise in most GNSS applications, since it can sometimes resemble other effects of interest (Braun et al., 2001; Larson et al., 2007), but in GNSS reflectometry it is instead the source of information.

2.1 Multipath signals in GNSS

The effect of multipath can be very varying as it depends on the topography and electromagnetic properties of the surroundings of the GNSS antenna (Nievinski and Larson, 2014a), but here we will focus on a narrow subset of multipath: signals that have been reflected only once and from a (mostly) flat, horizontal surface. Such multipath signals are highly coherent, which results in a coherent combination of the direct and the reflected signal, and it is this property that allows us to extract much information from the reflections.

In a coherent combination of two signals, the result depends on their phase offset, i.e. the delay between the two signals. From Figure 2.1 it is clear that both the reflector height $h$ (the height between the antenna and the reflecting surface) and the satellite elevation above the horizon $\varepsilon$ impact the path length difference between the direct and the reflected signals, and therefore the relative delay. As the surface acts as a mirror, the problem can be visualised as determining the difference in time of arrival for two receivers, one at the height $y = h$ above the surface, and and a virtual antenna below the surface at $y = -h$. 
Figure 2.1: The geometric delay of a GNSS signal reflected from a horizontal surface, such as the sea surface, depends on both the reflector height $h$ and the satellite elevation $\varepsilon$.

At low elevations, the path length difference will be small, going towards 0 at the horizon, and for zenith elevation the difference will reach $2h$. Because of the large distance to the satellites, the two incoming rays in Figure 2.1 can be approximated as being parallel (to within a milliarcsecond), therefore, the path length difference, or the geometric delay, can be written as (Georgiadou and Kleusberg, 1988)

$$\tau_G = 2h \sin(\varepsilon). \quad (2.1)$$

The composite power of the direct and the reflected signals, due to their interference, can be written as (Georgiadou and Kleusberg, 1988)

$$P = P_d + P_r + 2\sqrt{P_d P_r} \cos(\Phi), \quad (2.2)$$

where $P_d$ and $P_r$ are the power of the two signals, and $\Phi$ their relative phase offset, usually called interferometric phase in the field of GNSS reflectometry. Because of the elevation dependence of the geometric delay in Equation (2.1), the phase offset will vary across a satellite arc, and with it, the power of the composite signal.

In the receiver, it is the composite signal that is measured, and therefore its power $P$ defines the SNR that the receiver records, which thus also varies according to Equation (2.2). The cosine term in the equation creates a very distinct oscillatory behaviour in the signal strength as visible from Figure 2.2. These oscillations are the foundation of interferometric GNSS reflectometry. As we will see in the coming sections, their frequency, amplitude, and phase all contain information about the surface which the multipath signal was reflected off.

Signal strength measurements are usually stored in logarithmic units by GNSS receivers, and thus have to be converted to linear units in
Figure 2.2: Signal strength for the L5 signal from GPS satellite G10 recorded at the GTGU installation in Onsala, 25th of January, 2019. Left: $C/N_0$ in logarithmic units as recorded by the receiver. Right: SNR detrended with a 2nd order polynomial after first converting to linear units.

W/W before further processing. In the process it is important to note that while the signal strength measurements are referred to as SNR in GNSS-R literature, in practice, what is usually recorded in the receiver is the carrier-to-noise-density ratio ($C/N_0$) measured in dB-Hz. However, the two are closely related, and in logarithmic units their relationship can be written $C/N_0 = \text{SNR} + \text{BW}$, where only the receiver bandwidth BW has been added. Often, the manufacturers do not give all details about bandwidth and how they calculate $C/N_0$, and thus the conversion is mostly done assuming a bandwidth of 1 Hz (Larson et al., 2013b; Löfgren et al., 2014), i.e. assuming $C/N_0 = \text{SNR}$. While the true bandwidth for any realistic GNSS receiver is much higher (Joseph, 2010), the difference only amounts to a constant scaling factor in linear units. Therefore, under the 1 Hz bandwidth assumption, the SNR in W/W can be retrieved as

$$\text{SNR [W/W]} = 10^{C/N_0 [\text{dB-Hz}]/10}.$$  

The geometry shown in Figure 2.1 is somewhat simplified as it assumes optical rays with point-like reflections. Because of the frequencies used in GNSS reflectometry, the footprint of the reflections will be finite and cover a significant area. The size of the footprint is usually assumed to be roughly the first Fresnel zone (Larson, 2016; Larson and Nievinski, 2013), the size of which depends on the antenna height and the satellite elevation angle (Hristov, 2000, p. 323). While the footprints are easy to predict their considerable size at low elevation may create some less straightforward situations where part of the reflections come from for example the sea, and other parts from the shore. Thus, the signal may not always be as clear as the oscillations shown in Figure 2.2, which are for an almost optimal reflection geometry.
2.2 Lomb-Scargle retrievals

The most straightforward way of extracting information and measurements from the SNR oscillations is by spectral methods. By detrending SNR series with a low order polynomial, approximating the effect of $P_d + P_r$ in Equation (2.2), we are left with a roughly sinusoidal time series as seen in the right half of Figure 2.2, corresponding to the last part of Equation (2.2). The expression for the geometric delay in Equation (2.1) results in a variable interferometric phase

$$\Phi = \frac{4\pi h}{\lambda} \sin (\varepsilon) + \varphi,$$

(2.3)

which, when combined with the latter part of Equation (2.2), gives us an SNR residual which is proportional to

$$2\sqrt{P_d P_r} \cos \left(\frac{4\pi h}{\lambda} \sin (\varepsilon) + \varphi\right),$$

(2.4)

where $\lambda$ is the wavelength of the GNSS signal, and $\varphi$ is a residual phase delay. Thus, the SNR becomes a sine wave as a function of $\sin (\varepsilon)$ with a frequency of

$$f = \frac{1}{2\pi} \frac{\partial \Phi}{\partial \sin (\varepsilon)} = \frac{2h}{\lambda},$$

(2.5)

and therefore, by determining the frequency of the oscillations through spectral analysis it is possible to retrieve information about the reflector height and its changes. However, because of the non-regular sampling of the SNR data in $\sin (\varepsilon)$, ordinary FFT methods are inapplicable without resampling, and usually the Lomb-Scargle periodogram is used instead (Larson et al., 2009).

Each detrended SNR arc can be translated into a height measurement. As an example, the Lomb-Scargle spectra of the detrended SNR data in Figure 2.2 are presented in Figure 2.3. In this figure, the frequencies have been converted to their corresponding reflector height, $h = f \lambda/2$, showing a peak at just above four metre. Thus, by extracting the peak position we can retrieve one sea level measurement.

As seen in Figure 2.3, the spectra is often not a perfect peak. Sometimes other multipath sources influence the oscillation pattern, and create either a secondary peak or noise in the spectra. Therefore, it is common to screen the retrieved peaks with the condition that the peak power should be well above the noise floor. Larson et al. (2013a) suggest the criterion $P_{\text{peak}}/P_{\text{average}} > 3$, to remove uncertain measurements.

There are however a few caveats to translating the peak position to a reflector height measurement. The expression in Equation (2.5) assumes that both the residual phase $\varphi$ and the reflector height $h$ are constants, which may not necessarily be true. The effect and causes of varying
phase offsets will be discussed in Chapter 5, so for now we will focus on how to work around the assumption of a static reflector height. Early studies of interferometric GNSS reflectometry evaluated the acquisition of snow height (Larson et al., 2009) and soil moisture (Larson et al., 2008a). In both of these applications the assumption of a non-changing reflector height is either true, or approximately so, especially during the span of a single SNR arc, which spans roughly half an hour at most (Löfgren et al., 2014). However, when retrieving sea level with GNSS reflectometry, the assumption can no longer generally be said to be valid, since tidal changes can have a significant effect on the reflector height. Therefore, the retrieval method has to be corrected, and the expression for the frequency becomes

\[ f = \frac{1}{2\pi} \frac{\partial \Phi}{\partial \sin (\varepsilon)} = \frac{2h}{\lambda} + 2\sin (\varepsilon) \frac{\partial h}{\partial \sin (\varepsilon)} = \frac{2h}{\lambda} + \frac{2\dot{h} \tan (\varepsilon)}{\lambda \dot{\varepsilon}}. \]

Solving for \( h \) gives us

\[ h = \frac{f\lambda}{2} + \frac{\dot{h} \tan (\varepsilon)}{\dot{\varepsilon}}, \tag{2.6} \]

where the last part of the right hand side is the height rate correction term of Larson et al. (2013a).

One problem with the height formula in Equation (2.6) is obvious: the height is dependent on its own rate of change, which requires a time series of heights to estimate. As using height rates from a co-located tide gauge defeats the purpose of using GNSS reflectometry in the first place,

\[ ^1 \text{For the full derivation, see Appendix A.} \]
there are two different self-contained solutions. Either, the rate of change is estimated from tidal models, or the problem is solved iteratively by calculating approximate heights first, and then estimating the height rate to apply a correction. For locations where the sea level is dominated by tides, the former may work quite well, however, if the tides are small, and meteorological effects dominate the sea level variations, it will fail to realistically estimate the correction term. The latter method, meanwhile, is fully self contained and does not use any external information about the system that is studied. Therefore, that is the method of correction we chose for the studies in this thesis.

2.3 Lomb-Scargle in practice

Lomb-Scargle analysis is relatively easy to implement and use, and readily available software packages exist for anyone who wish to test GNSS reflectometry (Roesler and Larson, 2018). It is also straight forward to implement in most modern programming languages. The flow of such a program would be as follows:

1. Extract $C/N_0$ data from the RINEX files\(^2\), and convert from dB-Hz to linear units ($W/W$).

2. For each satellite arc, determine the azimuth and elevation time series using orbit data from either broadcast ephemerides, or from precise SP3 orbit files. The choice matters little as the small resulting differences in angles are negligible for GNSS-R retrievals.

3. Split the arcs in ascending and descending parts if applicable, and then detrend the SNR data with respect to elevation using low order polynomials.

4. Determine an angle mask to filter out SNR data from unwanted directions, i.e. with reflections from other surfaces.

5. Retrieve the spectral power densities as a function of $\sin(\varepsilon)$, using e.g. Lomb-Scargle analysis, for each of the remaining arcs. Then, assuming the expression of Equation (2.5) is valid, transform the peak positions to reflector heights.

6. Remove all peaks where the peak power is too low in comparison to the noise floor.

7. Use the retrieved reflector heights to estimate a height change rate, and calculate final heights according to Equation (2.6).

\(^2\)Tools for extracting SNR data are available at https://github.com/Ydmir/rinpy.
Chapter 3

New techniques

In the previous chapter, the focus was on how the oscillation frequency of SNR arcs corresponds to the vertical distance to the reflector. However, some features of the SNR oscillations depend on other properties of the surrounding of GNSS antennas (Nievinski and Larson, 2014a) and can therefore be used for environmental measurements. To retrieve parameters corresponding to these features we have to turn to other methods than spectral analysis, such as inverse modelling, which have previously been shown to be usable for measuring snow depth (Nievinski and Larson, 2014b,c).

In Section 3.1, I describe how to use least-squares adjustment and an inverse model to retrieve parameters from SNR data, especially in the case of sea surface measurements, based largely on Paper I. Then, in Section 3.2, I briefly explain how to use the inverse model in a Kalman filter to do GNSS reflectometry in realtime, as shown more detailed in Paper III. Finally, the last section of this chapter deals with an extension of the spectral analysis, by combining it with a particle filter, achieving close to realtime retrievals.

3.1 Inverse modelling of GNSS reflectometry

In inverse modelling, the desired property cannot be observed directly with the available equipment. Instead, some observable that is dependent on the desired quantity is observed, and by modelling the relation between the two, the desired property can be inferred. In inverse modelling of interferometric GNSS-R data, the observables are the SNR and the satellite elevation, and the desired quantities are the reflector height and other properties of the surroundings of the GNSS antenna.

When the model relating the observable to the desired quantity is known, the raw measurements can be translated into estimates of the properties, and if there is a one-to-one relationship between them, the process is straightforward. However, for a single GNSS SNR measurement
there is no unique height to which it corresponds, because of the oscillating nature of the SNR. Even if all other parameters (such as antenna properties, satellite signal strength, etc.) were known, a specific value for the SNR would still correspond to a large number of possible heights. Therefore, as in spectral GNSS-R methods, the analysis is performed using a longer series of detrended SNR data, and in order to retrieve physical properties from the data a model for the oscillations is fitted to the data using non-linear least-squares adjustment.

Model for SNR oscillations

The expression on which spectral analysis is based, given in Equation (2.4), is also the basis for the inverse modelling. From that expression we can see that the amplitude of the SNR oscillations depends on the direct and reflected power

\[ A' = \sqrt{P_d \cdot P_r}, \]

and while it is ignored in spectral analysis, both depend on the satellite elevation, making the oscillation amplitude \( A' \) to become a function of elevation. According to Nievinski and Larson (2014b), the direct power can approximately be expressed as

\[ P_d = P \cdot G_d, \]

where \( P \) is the incident RHCP power, and \( G_d \) the antenna gain in the satellite direction. Meanwhile, the reflected power becomes

\[ P_r = P|X|^2 S^2. \]

The factor \( X \) is a combination of both the antenna gain in the direction of the reflection point and the surface Fresnel coefficients, and accounts for the mix of RHCP and LHCP signal that arises in the reflection. The last factor, \( S \), arises from the loss of coherence of a signal from a reflection on rough surfaces (Beckmann and Spizzichino, 1987), and can be expressed as

\[ S = \exp \left( -\frac{k^2 s^2}{2} \sin^2(\varepsilon) \right), \]

where \( k = 2\pi/\lambda \) is the signal wavenumber, and \( s \) is a measure of the surface roughness.

In principle, it is possible to model everything but \( s \) with sufficient knowledge of the setup, including the knowledge of the satellite output power pattern as well as the surface composition. However, in practice both the exact satellite output power \( P \) and the gain pattern of the specific antenna \( G_d \), may often be unknown. Therefore, we choose a simplified model where the amplitude and phase of the oscillations are unknown parameters and all elevation-dependent amplitude effects are assumed to
New techniques

Figure 3.1: Example of fitting the model of Equation (3.1) to detrended SNR data.

be captured by the same exponential function. The model equation then becomes

$$\delta \text{SNR} = A \cdot \cos \left( \frac{4\pi h}{\lambda} \sin (\varepsilon) + \varphi \right) \cdot e^{-\Lambda k^2 \sin^2(\varepsilon)},$$  \hspace{1cm} (3.1)$$

where the unknown parameters are the maximum amplitude $A$, a phase offset $\varphi$, the combined antenna-surface damping coefficient $\Lambda$, and the reflector height $h$. If the antenna characteristics are known they can be included together with the Fresnel coefficients in the model function, thus decoupling $\Lambda$ from the gain pattern, making it rely only on the surface properties.

As the model has several parameters, but only one observable (i.e. the SNR), some assumption about the dynamics has to be imposed. Otherwise all parameters have to be estimated at each epoch, using only a single SNR measurement, making the solution underdetermined.

Amplitude, phase, and damping all depend on mostly antenna and surface properties which in most circumstances are quite stable, and therefore we choose to model them as being constant on the timescale of a day. However, the last unknown, height, is a completely different matter as it can vary on a much shorter time scale, when for example measuring sea level. Therefore, the height needs to be parametrized to allow for a time dependent height estimate, and for this our choice has fallen on using B-splines.

**B-Splines for modelling dynamic sea height**

B-splines were originally developed for computer graphics to approximate arbitrary curves and shapes with a finite set of parameters (Bartels et al., 1995). This makes them a prime candidate for parametrizing a variable reflector height for sea surface applications (Hobiger et al., 2014), as the
Figure 3.2: A B-spline consist of several overlapping basis functions.

sea surface dynamics can be very varying from site to site. Compared to for example using a tidal composition to represent the sea level, this allows for approximating sea surface heights even when wind and pressure are a dominant factor.

The B-spline is constructed recursively from the $0^{th}$ order basis function (Stollnitz et al., 1995)

$$N^0_j(t) = \begin{cases} 1 & \text{if } t_j \leq t < t_{j+1} \\ 0 & \text{otherwise} \end{cases}.$$ 

Higher order basis functions can then be defined as

$$N^r_j(t) = \frac{t-t_j}{t_{j+r}-t_j} N^{r-1}_j(t) + \frac{t_{j+r+1}-t}{t_{j+r+1}-t_{j+1}} N^{r-1}_{j+1}(t).$$

For a chosen degree $r$, this then allows us to approximate the sea surface height using a finite set of $N$ node heights $h_1, ..., h_N$ as

$$h(t) = \sum_{j=1}^{N} h_j N^r_j(t) + \Delta_{PCO}.$$ 

Here, the GNSS-frequency dependent vertical phase centre offset, $\Delta_{PCO}$, has been added to make the height refer to the antenna reference point regardless of which GNSS-signal frequency is used.

The time resolution of the B-splines is dependent on the number and density of the knots, $t_0, ..., t_{N+r+1}$, which define the boundaries of the nodes, where a denser placement of knots naturally leads to higher ability to resolve fast changes in sea surface height. Therefore, a high enough density of knots to resolve short time changes driven by wind and pressure changes is desirable. However, more knots also leads to more parameters in the least-squares analysis, with an increased computational burden as a consequence. Furthermore, past a certain number of nodes, increasing
the knot density further only increases the risk of overfitting, leading to a deterioration in the retrieval precision.

A major benefit of using a time-variable reflector height estimate is that there is no need to apply a height rate correction, in contrast to spectral analysis. Instead, the height change is already accounted for directly in the modelling. And whereas the correction term in Equation (2.6) is applied to a whole arc using the average values of elevation and height rate, using a time variable height instead uses the right values for each measurement epoch.

Implementing GNSS-R inverse modelling

The initial steps in retrieving information from SNR data using inverse modelling are very similar to the first steps in spectral analysis; the data has to be converted to linear units, be detrended to retrieve the oscillations, and screened to remove unwanted influences before further processing. However, unlike spectral analysis, the subsequent steps are not necessarily performed on each satellite arc independently. Instead all available data from a certain time span can be used, regardless of which satellite or frequency the signal comes from. Indeed, even satellites from different systems can be used in combination, rather than just performing the analysis on only GPS or Galileo etc. This is possible because the different signals share some information between them: most prominently the reflector height, but also, for example, the surface conditions. Other parameters are more dependent on GNSS-frequency and GNSS-system, and thus the parametrization has to reflect these different conditions. The number of parameters will therefore increase as follows:

**Amplitude:** Is partly dependent on the satellite transmission power, which varies with GNSS satellite system and generation (Steigenberger et al., 2018), but also for example on frequency-dependent reflection coefficients. Therefore the amplitude is estimated for each signal within each satellite block.

**Phase:** Depends mostly on antenna and surface properties which are GNSS-frequency dependent, which means that it must be estimated independently for each GNSS signal, i.e. GPS-L1, GPS-L2, GLONASS-L1, etc.

**Damping:** Varies mostly with surface properties and is therefore shared among all GNSS signals, satellites, and systems, so only one parameter is necessary.

**Height:** Also shared among all GNSS signals, satellites, and systems, but the number of parameters instead increases with the time over which the inversion is performed as more B-spline nodes have to be added.
Inverse modelling is an iterative process, where a set of parameters is tested to see how well they reproduce the output to be adjusted accordingly. Therefore, an initial guess for the state is needed to start the iterations. For most of the parameters, the choice of starting value does not affect the results very much, only slightly affecting the number of required iterations and thus computational load.

However, in the case of the height, the convergence is sensitive to the initial estimate and a careful choice has to be made. If the initial estimate and the true heights are too large apart the least-squares solver is prone to become stuck in a local optimum. This can be solved by for example increasing the search span of the parameters with global optimization techniques (Reinking, 2016), forcing the algorithm to search through a larger search space before narrowing down the final solution. However, we choose to tackle the problem differently, by improving the quality of the initial estimate, through using spectral analysis to find a rough estimate of the height. By first running a spectral analysis, we can use the resulting time series to fit an initial B-spline curve, which greatly increases the chance of convergence. This technique is especially helpful at GNSS-R sites where the tidal range is large, and any fixed initial height estimate would be too far off at least part of the time.

GNSS-R retrievals based on inverse modelling open up for several new possibilities in comparison to spectral methods. Firstly, the additional parameters that are retrieved mean that more information about the environment of the antenna can be inferred. One such result is presented in Chapter 4, where we use GNSS reflectometry to measure the presence of sea ice. Secondly, the model can be extended with additional details. The effects of antenna gain and surface reflection coefficients have already been briefly mentioned, and more effects will be discussed in Chapter 5. Furthermore, the inversion model does not necessarily have to be used with least-squares adjustment, but can also be the basis for other retrieval methods, as will be shown in the next section.

### 3.2 Kalman filtering for GNSS reflectometry

Kalman filtering is a type of sequential filtering technique, where the state of a system is predicted and updated each time new data is available. Compared to least-squares adjustment, which can often be applied to the same set of problems (Sorenson, 1970), the major benefit is the online processing which gives access to computationally efficient real time solutions.

I will only describe the Kalman filter in general terms here as it has been extensively studied and better explanations can be found elsewhere, e.g. (Brown and Hwang, 1992; Zarchan and Musoff, 2015). But in very brief terms, the basic idea behind the Kalman filter is to predict the state of a system, represented by the state vector $x$, and the corresponding
uncertainty, the covariance matrix $P$, given the best guess of the previous state of the system and a model of its dynamics, and then update the state estimate with new (noisy) measurements. In doing so, two models are needed: the dynamic model, explaining the time evolution of the system, and the measurement model, explaining the relationship between the state and the measurements.

In the original formulation, both of these relationships were assumed to be linear (Kalman, 1960), still the Kalman filter has found some of its greatest applications in nonlinear systems (Sorenson, 1970). Since the measurement model in GNSS reflectometry, i.e. Equation (3.1), is both sinusoidal and exponential, our problem clearly falls into this category. One way of tackling the non-linearity is with the extended Kalman filter (EKF), by linearisation of the problem (Jazwinski, 2007), however it is prone to divergence, especially when the system is highly nonlinear (Ljung, 1979). Another approach is the unscented Kalman filter (UKF) building on the unscented transform (Julier and Uhlmann, 1997), which has proven to perform well where the EKF fails (Wan and van der Merwe, 2000).

The unscented transform

The unscented transform is a method for determining the first moments of a random variable that has undergone a transformation. Where the EKF approach would be to locally linearise the transformation, which can often skew both mean and variance estimates (Julier and Uhlmann, 1997), the unscented approach is to propagate a minimal set of sigma points through the transformation, and using the transformed sigma points for estimating the statistics of the transformed variable (as illustrated in Figure 3.3). This has shown to provide more accurate estimations of the mean and covariance on many nonlinear systems than local linearisation can provide (Julier, 2002).

\[
\delta \text{SNR} = A \cdot \cos \left( \frac{\delta \phi}{\lambda} \sin(\varphi) \right) \cdot e^{-\Delta k^2 \sin^2(\varphi)}
\]

\[ y = f(x) \]

Figure 3.3: The unscented transform can estimate the mean and covariance of random variables undergoing even highly nonlinear transformations by propagating a minimal set of sigma points through the transformation.
In addition to its performance benefits, the lack of need to compute Jacobian matrices for the transformations can also make the algorithm easier to implement, as these can otherwise be non-trivial to derive (Julier and Uhlmann, 1997). This makes the UKF ideal for GNSS reflectometry, with its highly nonlinear measurement model.

**Dynamical models and B-splines in Kalman filtering**

In Kalman filtering, estimations of the temporal development of the parameters are needed, in addition to the measurement model. In the least-squares inversion of Section 3.1, the amplitude, phase, and damping were considered to be constant parameters. This works as the inversion has to be run on shorter batches of data, i.e. one or a couple of days, which allow the parameters to change on longer timescales. However, a Kalman filter could in principle be run indefinitely, just adding new measurements as they become available, and thus we have to allow the parameters to change.

All three parameters depend in part on surface conditions, and can therefore change with for example weather (Larson et al., 2008a) and growth season (Small et al., 2010). This makes their dynamic behaviour complex to explain and predict, so in the Kalman implementation we will treat them as random processes with small fluctuations, i.e. small process noise. This means that in the prediction step of the Kalman filter we will assume them to stay constant, but increase their corresponding uncertainty as they might have changed slightly in either direction, thus allowing the filter to change their values in the update step if it improves the residuals.

The functionality of the Kalman filter implementation depends on the chosen values for the process noises. If a process noise is too small, the filter will be unable to adapt to changes in the environment, loosing the ability to measure such changes and to provide more information than just reflector height estimates. However, if it is too large, the parameters can absorb some of the height variations, e.g. if the phase is allowed to vary quickly the filter may constantly vary the phase instead of setting the appropriate height, as these two effects would be indistinguishable in the model function, i.e. Equation (3.1). In our study in Paper III we found that the precision of the height estimate started to degrade if especially the phase and damping coefficient were given too large process noises.

If height is also modelled as a random process, similar concerns would apply. Too small process noise and the height filter would be unable to follow the natural variations caused by tides and meteorological phenomena, but if a too large process noise is used the filter may instead overfit and catch for example unmodelled elevation-dependent phase changes. As the true dynamics vary much between different sites, with varying tidal ranges and weather exposure, the right process noise can be hard
to define. Therefore, to sidestep this issue we continue on the same path as in the least-squares inversion, by using B-splines to represent the sea surface. These are modelled as constants, i.e. without process noise, but initiated with a high uncertainty so that they can be adjusted to fit the measurements.

As a set of B-spline nodes cover only a finite time, some scheme to renew their coverage is needed so that the state vector does not have to be excessively large, and the filter stopped and restarted each time the B-spline nodes run out. In our implementation, this is solved by keeping just the relevant node heights in the state vector using a special prediction step; once a B-spline node no longer has a bearing on a new measurement it is shifted out of the vector and a new node is added. Thus, the filter can run indefinitely while the state vector is kept to a minimum.

**Implementing realtime GNSS reflectometry**

As in both of the previously discussed algorithms, the Kalman filter uses the SNR oscillations as the basic measurement. Thus to really operate the filter in real time, the detrending of the SNR arcs has to also be performed in real time. In the other algorithms the detrending was done using data from a full satellite passage, thus removing the trend without removing the oscillations. However, in true realtime operation, the SNR data has to be detrended before a full satellite arc is available, and therefore another scheme is needed.

The trend of the arcs for a particular satellite are mostly constant, as it depends mostly on the antenna gain pattern and the satellite signal strength. Thus one solution is to include the arc trend in the model as a slowly varying process. However, even simpler to implement is using the average of the last few passages of the same satellite to fit the trend. This keeps the number of parameters in the state vector low, thus improving the computational efficiency and reducing the number of process noises that have to be determined for the filter to work correctly.

As shown in Figure 3.4, the UKF approach is only needed in the update phase, using the detrended SNR observations and an estimate of their noise levels, $R$. In the prediction step, the default prediction is that nothing has changed, which is trivially a linear prediction, and well suited for an ordinary Kalman filter approach. Only the covariance matrix $P$ is modified in the prediction, with the addition of the process noise covariance, $Q$. For simplicity and lack of better models, the process noise is assumed to be completely uncorrelated, i.e. that $Q$ is diagonal.

However, when a prediction step would take the filter out of the currently valid region of the B-spline approximation, a new B-spline needs to be introduced and a special prediction step is applied. First a normal prediction step is applied, i.e. modifying the uncertainties, and then the shift operation is applied. All the node height estimates in the state vector
are shifted by one position, which means that one node is shifted out of the vector, and a new one introduced. Without adding a tidal harmonic analysis to the model, the best estimate for the new node is the same value as the node immediately before it. In addition, the covariance matrix has to be changed accordingly, shifting the rows and columns corresponding to the height nodes. The row and column of the new node are again copied from the previous node, but with increased self covariance to signify the uncertainty of the height assumption.

Figure 3.4: Visualisation of the Kalman filter algorithm, using the unscented transform to perform the update step while still using the normal prediction step. However, a special shift step is introduced each time a new B-spline node is needed to cover the new time step.

3.3 Particle filter with Lomb-Scargle periodograms

The final method presented in this thesis is an extension of the spectral methods presented in the previous chapter. Instead of just using the peak position of a Lomb-Scargle spectra as a height measurement, the whole spectra can be interpreted as a probability distribution for the position of the reflector height. Here, we will use this in a particle filter, as the measurement probability, to combine all Lomb-Scargle spectra into a time series of reflector heights.

Particle filters, like Kalman filters, belong to the family of Bayesian filters, but are inherently more suited to model nonlinear and multimodal problems (Fox et al., 2003). While a standard Kalman filter models a linear process using only the mean and the variance to represent process and measurement noise distributions, thus only giving exact solutions
for Gaussian distributions, particle filters use an ensemble of particles to estimate the distributions, giving an approximate solution for any kind of distribution (Arulampalam et al., 2002).

In Bayesian filtering, there are two distributions that define the dynamics: the probability of measurement $x_{\tau}$ given the previous state $x_{\tau-1}$, $p(x_{\tau}|x_{\tau-1})$, and the observation likelihood, $p(z_{\tau}|x_{\tau})$, that the measurement $z_{\tau}$ comes from the state $x_{\tau}$. It is for the second distribution that we will use the Lomb-Scargle spectra, and thus interpret it as the likelihood of the spectra arising from a particular sea state.

Like the Kalman filter, the particle filter also includes a model of the dynamics of the state. To keep the complexity low and to keep it simple to implement, we choose to use a two variable state $x_{\tau} = (h_{\tau}, v_{\tau})$, where $h_{\tau}$ is the reflector height, and $v_{\tau}$ the height change rate, thus $x_{\tau+1} = x_{\tau} + v_{\tau} \cdot \Delta t$.

The uncertainties of the state model, i.e. the system noise, is simulated by adding random noise to the height and its rate of change. A bonus of including the height rate in the model is that the height rate correction of Equation (2.6) can be included directly, without iterating.

After retrieving Lomb-Scargle spectra following the outline in Section 2.3, the particle filter algorithm can be described by the following steps, illustrated in Figure 3.5.

1. Initiate $N$ particles, $x^{i}_{0} = (h^{i}, v^{i})$, $i = 1, ..., N$, and give them equal weights, $w^{i}_{0} = \frac{1}{N}$.

2. Propagate states and add random noise $W$ and $V$ for height and change rate, respectively:

$$h^{i}_{\tau} = h^{i}_{\tau-1} + v^{i}_{\tau-1} \cdot \Delta t + W,$$
$$v^{i}_{\tau} = v^{i}_{\tau-1} + V,$$

where $\Delta t = t_{\tau} - t_{\tau-1}$.

3. Compute the probability that $x^{i}_{\tau} = (h^{i}_{\tau}, v^{i}_{\tau})$ gave rise to the measurement, i.e. the Lomb-Scargle spectra, and scale the weights accordingly:

$$w^{i}_{\tau} = w^{i}_{\tau-1} \cdot P_{\text{lomb}}(h^{i}_{\tau}) \cdot p(v^{i}_{\tau}),$$

where $p(v)$ is the probability of having a change rate of $v$. Then renormalize the weights:

$$w^{i}_{\tau} = \frac{w^{i}_{\tau}}{\sum_{j=1}^{N} w^{j}_{\tau}}.$$ 

4. Resample $N$ new particles based on the weights $w^{i}_{\tau}$ using an appropriate resampling strategy (Hol et al., 2006), and reset the weights to $w^{i}_{\tau} = \frac{1}{N}$. 
5. Calculate the estimated value of the state vector, $\hat{x}_\tau$.

6. Step up the time index, $\tau = \tau + 1$, and go to step 2.

The extra scaling of $p(v)$ in step 3 restrains the change rate to realistic values, and is implemented as a Gaussian distribution centred around $v = 0$. Similar results can be achieved by another implementation of the random noise instead. The implementation thus contains three parameters that have to be determined for a functioning filter: the process noise of the height and its rate of change, as well as the width of the distribution $p(v)$.

A short test campaign indicates that the particle filter works on par with smoothed Lomb-Scargle retrievals, as seen in Figure 3.6. Both the particle filter and the Lomb-Scargle algorithm, using a moving average of 50 retrievals, produced a RMS error of 2.6 cm. However, while such averaging of Lomb-Scargle results depends on data recorded both before and after the time in question, the particle filter only depends on the past information. Thus, similar to the motivation of using Kalman filtering over least-squares inversion, particle filter could see a possible usage in applications where near-realtime results are desirable.

Figure 3.5: The Lomb-Scargle spectra can be used as the measurement probability in a particle filter. Particles corresponding to high power in the spectra are given higher weight. The particle filter can easily handle possible multi-modality of the spectra, as it is approximated by an ensemble of particles.
Figure 3.6: **Top:** Time series of all Lomb-Scargle spectra from 3 days overlaid with particle filter particles and maximum likelihood estimates for each epoch. The particle marker sizes are scaled with the particle weights. **Bottom:** Sea level measurements from the particle filter compared to results from Lomb-Scargle retrievals as well as the tide gauge records. The smoothed line uses a running average of 50 retrievals. While the results in the bottom panel were calculated with 2000 particles to achieve high accuracy, the top panel includes only 40 particles at each epoch for clarity.
Finally, comparing to the Kalman filter implementation of Paper III, the particle filter has much fewer parameters, making it simpler to implement and adjust for a particular installation. Furthermore, being based on the Lomb-Scargle retrievals it is computationally less complex. Thus, if a simple, low latency implementation is wanted, particle filtering may just be the way.
Chapter 4

New applications

Since its inception, GNSS reflectometry has found several unanticipated uses (Larson, 2019), and with new methods and new technology come even more opportunities for environmental measurements with reflected GNSS signals. In this chapter I briefly introduce how to both measure new properties of the environment, and how to make GNSS reflectometry mobile, based on Papers II and IV.

4.1 Detecting sea ice

An improvement of both the Kalman implementation and the least-squares inverse modelling is that they retrieve several parameters describing the shape of the SNR oscillations, not only the frequency or the reflector height. These parameters, as explained previously, depend partly on properties of the reflecting surface, especially the damping parameter. The damping parameter is affected by both the electromagnetic properties of the reflecting surface and its roughness, and therefore becomes sensitive to the exact surface conditions.

In the transition from open water to ice, both the physical appearance of the sea surface and its electromagnetic properties can change (Eicken, 2003), which should make the damping parameter sensitive to this transition. That GNSS reflectometry can indeed be used to detect sea ice is shown in Figure 4.1, where it can be seen that the damping significantly changes during the period in which the Swedish Meteorological and Hydrological Institute reported sea ice in the vicinity of the GTGU installation at Onsala Space Observatory.

As information about the sea ice coverage is of great importance for instance for marine shipping (Jevrejeva and Leppäranta, 2002), and sea ice formation usually starts at the land-water boundary (Granskog et al., 2006) where the low resolution of spaceborne remote sensing techniques makes sea ice detection difficult, a network of ground-based, coastal GNSS receivers could perceivably be used in conjunction with other techniques to
Figure 4.1: **Left:** Time series of the retrieved damping parameter using Kalman filtering and least-squares inversion at GTGU. The dark grey dashed lines show the start and the end of the period in which SMHI reported sea ice in the vicinity of the installation. The light grey dotted line shows the date of the photo on the right. **Right:** Photo of GTGU taken on 6th of February, 2012. Photo: Johan Löfgren.

improve the detection of coastal sea ice growth. The same installations can naturally also be used as tide gauges – as shown in the previous chapters – and still provide data for tectonic studies, which shows the versatility of the GNSS-R technique and its capacity to combine several measurements in one sensor.

### 4.2 Mobile GNSS reflectometry

In parallel with the development of more advanced techniques for GNSS-R retrievals, the underlying hardware has undergone improvements as well. There are now more GNSS satellites than ever, and many of them transmit more signals in addition. Currently there are over 100 operational GNSS satellites in orbit – when combining all four constellations, GPS, GLONASS, Galileo, and BeiDou – and it is not uncommon to have 50 satellites or more in view at any given time and location.

At the same time the GNSS receivers have been developed to utilize all new available signals, and high-end receivers can track several hundred signals at once, with sampling times well under a second. But it is not only the high end of the receiver market that has seen improvement. GNSS chips for everyday devices have also improved, with more capacity and precision.

Modern mobile telephones and tablet computers now carry GNSS chips capable of recording GNSS signals from all available systems, albeit mostly with limited receiver bandwidth and thus only recording signals in the L1/L5 frequency band. Furthermore, with the increased computational
capacity of the phones and tablets, some new mobile devices are even capable of recording raw GNSS signals and make these available for post processing. Consequently, smart devices are currently capable of precise network positioning (Realini et al., 2017).

The same recorded data that makes network positioning possible on a mobile device is also the data that is necessary for GNSS reflectometry. Thus, GNSS reflectometry is feasible using such devices, and as shown in Figure 4.2 (with more details in Paper IV), the tablet seems to perform on a similar level as the high-end equipment of the GTGU installation. While the analysis was done on a separate computer in this test implementation, the computational capacity of a mobile phone or tablet computer allows for implementing the full process on the device itself. Thus, mobile GNSS reflectometry could become a viable low budget alternative to tide gauges, but also for snow measurement, soil moisture sensing, and vegetation monitoring, where simple GNSS reflectometry algorithms have proven useful (Larson, 2019). This could lead to a more widespread usage of the technique, and make it available to a wider public.
Chapter 5

Atmospheric delay and absolute leveling

A radio signal propagating in the atmosphere will be affected and have a different propagation path than what would be experienced in vacuum. The topic has been well studied, as it is important in for example radio astronomy imaging (Rohlfs and Wilson, 2013), and for positioning applications in GNSS (Hofmann-Wellenhof et al., 2007). The effect has also been shown to be of importance in GNSS reflectometry (Santamaría-Gómez and Watson, 2017; Williams and Nievinski, 2017), where it can significantly affect the mean retrieved reflector height. But how to properly account for it is still a topic of research. In this chapter we will look at both already suggested correction methods, as well as new ones, for spectral analysis as well as inverse modelling methods.

5.1 Effect of the atmosphere

There are two main effects of the atmosphere on the propagation of radio waves: the variable index of refraction causes the signal path to deviate from a straight line, and the speed of propagation is lower in the atmosphere than in vacuum. These effects, the refraction and the delay, are illustrated in Figure 5.1. While the two effects have the same source, they are quite different in nature, and as a first approximation they can be treated separately. In GNSS reflectometry they will mainly affect the height retrieval, as they affect the oscillation frequency of the detrended SNR.

The first effect to be included in GNSS reflectometry was the atmospheric refraction, accounting for the changed angle of incidence that is caused by the curved propagation path, leading to a higher apparent elevation than vacuum elevation (see Figure 5.1a). The refraction is straight-forward to account for — both in methods based on inverse modeling as well as spectral analysis — as the true incidence elevation,
Atmospheric refraction and delay affect GNSS reflectometry.

Firstly, the signal will be refracted and thus have a higher true elevation than the vacuum elevation to the satellite. Secondly, the signal propagation speed will be slowed down slightly by the atmosphere. The latter will affect the reflected signal slightly more than the direct signal, leading to additional phase delays.

Also called apparent elevation, can be retrieved by adding a correction term to the vacuum elevation angle (Santamaría-Gómez and Watson, 2017),

$$\varepsilon^a = \varepsilon + \delta \varepsilon.$$ 

In the work by Santamaría-Gómez and Watson (2017), the approximation formula for the refraction correction given by Bennett (1982) was used, which, when modified for change of units, becomes

$$\delta \varepsilon = \frac{1}{60} T + \frac{283}{273} \frac{P}{1010.16} \cot \left( \frac{\varepsilon}{\varepsilon + 4.4} \right),$$

where the temperature $T$ is given in degree Celsius, the pressure $P$ is given in hectopascal, and the elevations are expressed in degree.

The correction term is elevation dependent, with a larger effect at low elevations (Bennett, 1982). Therefore, at the antenna, the vacuum elevation span travelled by a satellite according to the orbit files will be larger than what is experienced at the antenna. Thus, if the vacuum elevations are used to retrieve the SNR oscillation frequency, the frequency will be slightly underestimated, leading to a too low reflector height, or conversely a too high sea level.

The second effect is the lowering of the speed of propagation of the signal, caused by the troposphere. As this effect affects the reflected signal slightly more – it passes through a lower part of the atmosphere with
slower propagation speed – it will lead to an additional interferometric delay. Thus, the geometric delay of Equation (2.1) has to be complemented with a tropospheric component to form a total delay

\[ \tau = \tau_G + \tau_T = 2h \sin(e^a) + \tau_T. \]

According to Williams and Nievinski (2017), the tropospheric delay component can be written as

\[ \tau_T = 2\Delta \tau^z_h \cdot m_h(\varepsilon) + 2\Delta \tau^z_w \cdot m_w(\varepsilon), \]

where \( \Delta \tau^z_X = \tau^z_X(-h) - \tau^z_X(0) \) are the differences in the wet and hydrostatic zenith tropospheric delay components, and \( m_X(\varepsilon) \) the respective mapping functions. Following the work of Williams and Nievinski (2017) the VMF1 mapping functions (Böhm et al., 2006) are used, and the delays are calculated using the Global Temperature and Pressure (GPT2w) model (Böhm et al., 2006), so that the corrections can be calculated at any station regardless of access to real weather data. Note that we here use the vacuum elevation to be consistent with the formulation of the mapping functions.

The difficulty of correcting for the effect of the troposphere depends on which retrieval method that is used. In the methods based on inverse modelling, it is relatively straightforward as it will amount to an extra phase delay \( \Delta \varphi = k\tau_T \), which can be added directly to the model function in Equation (3.1). For spectral retrievals a correction method has to be employed instead, as it cannot directly be included in the retrieval procedure. Williams and Nievinski (2017) suggests that a correction term, similar to the height rate correction of Equation (2.6) can be employed, noting that

\[ f \cdot \lambda = \frac{\partial \tau}{\partial \sin(e^a)} = 2h + 2h \frac{\tan(e^a)}{e^a} + \frac{\partial \tau_T}{\partial \sin(e^a)}, \quad (5.1) \]

which gives a tropospheric height correction

\[ \delta h_T = -\frac{1}{2} \frac{\partial \tau_T}{\partial \sin(e^a)}, \quad (5.2) \]

which has to be added to the retrieved reflector height from Chapter 2. However, there are a few caveats to using this correction term. As the authors note in introducing the correction term, it is an instantaneous correction taken at a specific value of the elevation, the choice of which will necessarily affect the correction term. In principle this means evaluating the correction term at the average elevation of the arc. However, unless the elevation range is small, the bias will change significantly during the arc. The effect is non-linear, with exponentially larger offsets at low elevations, and thus its average will not correspond to the value of
the average elevation of the arc. Furthermore, the amplitude of the SNR oscillations decrease with higher elevation, thus low elevations are weighted more in the spectrum analysis. Thus, the average elevation is not the optimal choice for evaluating Equation (5.2), and to more accurately assess the magnitude of the correction, the equation has to be integrated over the full elevation range, using weights corresponding to the decreasing oscillation amplitudes. This is not easily done in spectral analysis, as several parameters of the shape of the oscillations are unknown.

As an alternative, the input to the inversion routine has to be adjusted to mitigate the effect of the troposphere. In inverse modelling methods this was achieved by adding the delay to the model. In spectral methods, we can instead perform a variable substitution, similar to how Santamaría-Gómez and Watson (2017) remove the effect of sea surface variations, i.e. exchanging $\varepsilon$ for $\varepsilon'$, such that

$$f \cdot \lambda = \frac{\partial \tau}{\partial \sin (\varepsilon')} = 2h.$$

Exchanging $2h$ in the right hand term of Equation (5.1), and assuming that $\dot{h} = 0$, we can in a finite approximation arrive at

$$\Delta \sin (\varepsilon') = \frac{1}{1 - \frac{\Delta \tau}{\Delta \tau}} \Delta \sin (\varepsilon^a). \quad (5.3)$$

Thus, the troposphere can be accounted for through exchanging $\sin(\varepsilon^a)$ by

$$\sin(\varepsilon'_n) = \sin(\varepsilon^a_0) + \sum_i^n \Delta \sin (\varepsilon'_i)$$

when calculating the spectral densities. Note however that the tropospheric delay component is also height dependent, thus to calculate the correction term in Eq. (5.3), an estimate of the height is needed. To solve this, we calculate the correction in an iterative fashion, first estimating the height ignoring the tropospheric effect, then calculating a modified sine of elevation as the input to the next iteration of the spectral analysis.

### 5.2 Applying the corrections

Previous GNSS-R studies have only used one of the two aspects of the atmosphere, either the atmospheric bending (Santamaría-Gómez and Watson, 2017) or the tropospheric delay (Williams and Nievinski, 2017). However, as noted in the previous section, the effect of the two are of entirely different natures, where the first one changes the geometry and the other slows down the signal. Therefore, to fully correct for the atmospheric influence, both of these effects have to be corrected for. Ideally,

\[1\] The full derivation can be found in Appendix A.
their combined effect should be evaluated using ray tracing techniques (Nikolaoud et al., 2020), but as the computational costs of employing ray tracing in the GNSS-R retrievals would be prohibiting, we use the above mentioned approximations.

The results of applying the different corrections for both inverse modelling and Lomb-Scargle retrievals are shown in Figure 5.2, where the difference between the sea level measured by the tide gauge and the GNSS-R installation at Onsala is shown. From the figure it can be seen that while the sea level results are shifted by the same distance for both GNSS-R algorithms when only correcting for atmospheric refraction, the picture looks different when correcting for the tropospheric delay. In the case of atmospheric refraction, the correction is handled exactly the same in both algorithms – correcting the elevation on the ‘observation’ level – thus it is expected that their average sea level estimation should coincide. However, when addressing tropospheric delay, the exact calculated delay is added directly to the inversion model whereas it is only applied as an
approximate correction in the case of Lomb-Scargle analysis. In Figure 5.2, the difference in handling is visible as an offset between the sea level measurements.

The correction term of Williams and Nieviski (2017) is seen to underestimate the correction compared to inverse modelling with tropospheric delay, by approximately 35%. The mean elevation of the arcs are on average 6.0°, which corresponds to an offset of roughly 9 cm when using Equation (5.2), which is seen as the distance between the lines b and c in Figure 5.2. However, as discussed above, the offset rather corresponds to using Equation (5.2) for a lower elevation, where the 14 cm offset seen between inverse modelling with and without troposphere delay modelling corresponds to the value of the correction term at 4.5°. In comparison, the newly suggested correction method corrects for 90% of the average offset between line II and III in Figure 5.2, showing that our suggested variable substitution is better at correcting for the tropospheric delays. However, as it is still only an approximate correction, its effect is still smaller than when adding the modelled delay directly to the retrievals, as done in inverse modelling.

5.3 Absolute leveling of GNSS reflectometry

One of the proposed advantages of GNSS reflectometry is the inherent capacity of correcting for vertical land motion and tying the measured sea level to an absolute reference point (Nieviski et al., 2020). However, in most GNSS-R studies to date, only relative sea level is measured, and the few that do try to measure absolute sea level often find significant offsets in at least some of the measurements (Santamaría-Gómez and Watson, 2017; Williams and Nieviski, 2017).

The atmospheric effects have been shown to be one of the contributions to the vertical offset of the GNSS-R retrievals, however Williams and Nieviski (2017) found that while they underestimated the reflector height without tropospheric corrections, they overestimated it when applying their correction method. As shown in the previous section, the results would most likely overshoot even more when applying a more rigorous correction. Furthermore, in the study they neglected the correction for atmospheric refraction which would have increased the reflector height even further. This is consistent with the findings in Figure 5.2, where the reflector height is seen to be overestimated with roughly 20 cm when accounting for both atmospheric refraction and tropospheric delay.

Even though the atmospheric corrections make the reflector height estimate overshoot, the underlying physical principles are well understood and documented, and therefore they should be included in GNSS reflectometry for correctness. It can also be seen in results that the corrections are necessary. For example, Williams and Nieviski (2017) found that, by applying their tropospheric correction term, they reduced the elevation de-
Figure 5.3: RMS error of GNSS-R retrievals from GTGU using inverse modelling compared to the Onsala tide gauge, with (colour) and without (greyscale) tropospheric delay modelling.

Pendency of residuals. And, in inverse modelling, incorporating the model for tropospheric delay reduces the RMS error of the retrieved sea level, as presented in Figure 5.3. While the effect is small, it is still significant and consistent for all satellite systems, showing that the correction indeed removes systematic effects. Therefore, the offset must be explained by some other, yet to be corrected for, phenomena.

Candidates for such phenomena must fulfil two criteria: the effect must differ for the direct and the reflected signal, and it must be elevation dependent. If it fails to meet either, there will be no effect on the reflector height. Unless there are differences between the effect on the direct and the reflected signal, there will be no change in the interferometric phase. This excludes for example satellite errors and ionosphere effects. And if there is no elevation dependence, the effect will just be folded into the constant offset in Equation (2.3), giving no effect on the oscillation frequency.

One candidate phenomena is the electromagnetic bias; as wave throughs are wider than the peaks, more energy is scattered at the bottom than at the top, shifting the average reflection point downwards (Ghavidel et al., 2016). However, in GNSS-R studies this has been found to be a minor effect, with a maximum of only 4 cm at a site that is more exposed to wind and waves than GTGU (Sun, 2017). Wind-driven offsets have been observed at Onsala as well, but the offset only reaches significant values at extreme wind events (Nieviniski et al., 2020), thus it can only explain a minor part of the 20 cm offset seen in Figure 5.2.

A more likely candidate is the combined interaction of the Fresnel reflection coefficients and the antenna pattern. In the reflection, the phase
of the circularly polarized GNSS signals are shifted with the argument of the Fresnel coefficient, which depends on the angle of incidence (Hristov, 2000). Furthermore, some power is shifted between the two circular polarisations. The effect is small at grazing angles, but almost complete at zenith where the reflected signal will be mostly LHCP. Since a GNSS antenna is not completely insensitive to LHCP signals this will lead to a combination of the two components in the receiver, and the interferometric phase will depend on the composite signal. According to Nievinski (2020), the effect can be on the order of 10 cm, depending on elevation, signal, and antenna.

To test the effect of the coupled antenna-surface reflection coefficients, we used the multipath simulation software MPSIM (Nievinski and Larson, 2014a) to calculate the induced phase shift as a function of elevation. The calculated phase shift was then added directly into the model function in Equation (3.1), to remove its effect. The result is an upward shift of 6 cm, still short of explaining the remaining residual. However, only the gain of the antenna pattern was available, and thus effects arising from the phase center variation were neglected. To fully model the coupling, the full antenna pattern is needed, including phase center variations for both RHCP and LHCP, and both negative and positive elevations.
Chapter 6

Summary of appended papers

Paper I

**Improving GNSS-R sea level determination through inverse modeling of SNR data**
In this first paper we developed and tested a new method to retrieve environmental parameters from GNSS SNR data. The method is based upon inverse modelling using least-squares adjustment as opposed to Lomb-Scargle analysis, which has been the prevalent method. This development resulted in more precise measurements, partly due the method being able to use data from several satellites at once.

For this article, I was responsible for developing the idea and running the analysis. I then prepared the manuscript in collaboration with my co-workers.

Paper II

**Coastal sea ice detection using ground-based GNSS-R**
An addition, and improvement, of the inverse modelling is that it can retrieve several different properties of the GNSS SNR data. Some of the parameters in the inverse retrievals are sensitive to the electromagnetic properties of the reflecting surface. For a sea surface, this results in a detectable signal when the sea transitions from open to frozen. Our work in this paper resulted in the first demonstration of coastal GNSS-R being sensitive to sea ice conditions.

My contributions to this article consisted of making the data analysis as well as preparing and being main responsible for the manuscript.
Paper III

Real-time sea-level monitoring using Kalman filtering of GNSS-R data

Based on the inversion model used in the previous two papers, we developed and demonstrated a method to retrieve measurements in real time using Kalman filtering. The results proved to be on a similar level as the results from least-squares inversion, making real time GNSS-R feasible.

In the preparations for this article I developed and tested an implementation of the method and ran the analysis. I prepared the manuscript, which was then finalized together with my co-authors.

Paper IV

Can we measure sea level with a tablet computer?

For the fourth paper, we decided to test the capacity of modern phones and tablet computers to perform environmental measurements. Newer models of mobile phones and tablets are able of recording raw GNSS data. Thus it is possible to do GNSS reflectometry with handheld devices. We demonstrated that the precision of sea level retrievals using GNSS-R was on the same level of precision when using a mobile device as when using high-end geodetic equipment. This opens up for mobile and inexpensive GNSS-R measurement campaigns.

For this paper I proposed and tested the idea. I also wrote the manuscript together with my co-author.
Appendix A

Derivations

Height rate correction

The height rate correction is a result of the height being time dependent. Starting from the expression for the geometric phase delay

\[ \Phi = \frac{4\pi h}{\lambda} \sin(\varepsilon) + \varphi, \quad (A.1) \]

we can compute the instantaneous frequency in \( \sin(\varepsilon) \) as

\[ f = \frac{1}{2\pi} \frac{\partial \Phi}{\partial \sin(\varepsilon)} = \frac{2h}{\lambda} + 2 \sin(\varepsilon) \frac{\partial h}{\partial \sin(\varepsilon)}. \quad (A.2) \]

The reflector height is not directly dependent on the satellite elevation, however as they both vary with time we can reformulate the derivative of the height as

\[ \frac{\partial h}{\partial \sin(\varepsilon)} = \frac{\partial h}{\partial t} \frac{\partial t}{\partial \sin(\varepsilon)}. \quad (A.3) \]

The first factor on the right hand side is simply the rate of change of the height \( \dot{h} \), and the second can be calculated as

\[ \left( \frac{\partial \sin(\varepsilon)}{\partial t} \right)^{-1} = \frac{1}{\dot{\varepsilon} \cos(\varepsilon)}. \quad (A.4) \]

Combining the above equations results in

\[ f = \frac{2h}{\lambda} + \frac{2\dot{h} \sin(\varepsilon)}{\lambda \dot{\varepsilon} \cos(\varepsilon)} = \frac{2h}{\lambda} + \frac{2 \dot{h} \tan(\varepsilon)}{\lambda \dot{\varepsilon}}. \quad (A.5) \]

Finally, solving for \( h \) we have

\[ h = \frac{\lambda f}{2} - \frac{\dot{h} \tan(\varepsilon)}{\dot{\varepsilon}}, \quad (A.6) \]
where the last part is the height rate correction. While the rate of change for the satellite elevation, $\dot{\epsilon}$, as the elevation itself be computed from orbit files, the rate of change for the height is an unknown in the retrievals. One way of retrieving $\dot{h}$ is to first retrieve reflector heights from the Lomb-Scargle spectra under the assumption $\dot{h} = 0$, and using the retrieved time series to calculate a better estimate. This new estimate is then used to correct the retrieved heights.

**A new correction for troposphere**

With the addition of troposphere, the expression for the phase delay in Equation (2.3) instead becomes

$$\Phi = 4\pi h \frac{\sin(\epsilon)}{\lambda} + \Phi_T + \varphi,$$

where

$$\Phi_T = \frac{2\pi}{\lambda} \tau_T$$

is the additional phase delay introduced because of the tropospheric path delay $\tau_T$.

Now, analogous to Equation (2.5), the instantaneous frequency instead becomes

$$f = \frac{1}{2\pi} \frac{\partial \Phi}{\partial \sin(\epsilon)} = \frac{2h}{\lambda} + \frac{1}{2\pi} \frac{\partial \Phi_T}{\partial \sin(\epsilon)}.$$

Williams and Nievinski (2017) recognised the last part as a correction term that can be added after the reflector heights have been retrieved, similar to the height rate correction in the previous section. However, as explained in Chapter 5, there are problems with choosing which angle to evaluate the correction at, resulting in poor accuracy in the correction.

Instead, we recognise that we can perform a variable substitution from $\epsilon$ to $\epsilon'$, or rather $\sin(\epsilon)$ to $\sin(\epsilon')$, such that

$$f' = \frac{1}{2\pi} \frac{\partial \Phi}{\partial \sin(\epsilon')} = \frac{2h}{\lambda}.$$

Combining Equations (A.9) and (A.10) results in

$$\frac{\partial \Phi}{\partial \sin(\epsilon)} - \frac{\partial \Phi_T}{\partial \sin(\epsilon)} = \frac{\partial \Phi}{\partial \sin(\epsilon')},$$

or equivalently

$$\frac{\partial (\Phi - \Phi_T)}{\partial \sin(\epsilon)} = \frac{\partial \Phi}{\partial \sin(\epsilon')}.$$

From this we can develop an expression of how $\sin(\epsilon')$ changes with $\sin(\epsilon)$:

$$\frac{\partial \sin(\epsilon')}{\partial \sin(\epsilon)} = \frac{1}{1 - \frac{\partial \Phi_T}{\partial \Phi}}.$$
Going back to path delays instead of phase delays, and approximating the expression to finite differences, we finally have

\[
\Delta \sin (\varepsilon') = \frac{1}{1 - \frac{\Delta \tau}{\Delta \tau}} \Delta \sin (\varepsilon).
\]  

(A.14)

Therefore, by starting at an initial elevation \(\varepsilon'_0 = \varepsilon_0\), the adjusted sine of elevation can then be calculated sequentially as

\[
\sin(\varepsilon'_n) = \sin(\varepsilon_0) + \sum_{i}^{n} \Delta \sin (\varepsilon'_i), \quad n \neq 0.
\]  

(A.15)

To retrieve the corrected height estimate, these new elevations are used in the Lomb-Scargle retrievals, producing a frequency \(f'\) which can be used in e.g. Equation (A.6) instead of the true oscillation frequency \(f\).
Bibliography


Joseph, A. (2010). What is the difference between SNR and $C/N_0$. *Inside GNSS*.


